Álvaro Corrales Cano

Data Scientist and Economist with 4 years of experience in data analysis in client-facing roles in tech and consulting firms

📱 +44 7395 620201 　　　 ✉ aldcorrales@outlook.com 　　　 📍 London, UK

in linkedin.com/in/alvaro-corrales-cano 　　　 🔗 alvarocorrales.github.io

Technical skills

Methods

Machine Learning experience
- Natural Language Processing (NLP) – Sentiment Analysis, Topic modelling (LDA), Embeddings (word2vec, doc2vec)
- Linear and logistic regression, SVM, KNN, Random Forest, Gradient Boosting
- Clustering methods – Hierarchical, K-means

Econometrics
- Causal inference – Panel data, Randomised Controlled Trials
- Discrete choice – Classification problems
- Time series – ARIMA, GARCH
- Duration Analysis

Other
- Deep Learning – LSTM
- Bayesian Networks

Technologies

Python – main programming language
- Data wrangling: *numpy, pandas, scipy*
- ML: *scikit-learn, statsmodels, keras, xgboost, lightgbm*
- Visualisation & dashboarding: *matplotlib, seaborn, plotly, streamlit*
- NLP: *gensim, nltk, spacy*

Cloud
- IBM Cloud and Cloud Pak for Data
- Docker

Other
- R, Spark, SQL, Stata, Excel

Professional experience

IBM
Data Scientist

London, UK
November 2019 – Present

IBM Garage - Helping organisations use their data with **Cloud** technology, **AI** and an **Agile** methodologies. Show the value of IBM's Cloud Pak for Data through PoC and MVP. Relevant experience:

- **Emergent Alliance** - **Led two teams** producing open-source material in the fight against Covid-19. Code available on Github.
  - i. Created a model of the economy to study how economic shocks propagate across economic sectors. We spined it of as a publicly available app.
  - ii. Identified mental health problems and perception of vaccines for NHS on Twitter using topic modelling (LDA), sentiment analysis (Watson NLU) and LSTM.
- **Publishing industry** – Created a ML pipeline to identify authors of comments on various online forums using doc2vec and K-means clustering.
- **Energy sector** – Improved the data collection and organisation process of a company. Created a harmonised catalogue with data protection rules, which can be partially accessed by external users, as required by national regulations.

Other

- o I have **spoken in public events** with organisations such as Code First Girls and PyData London, for which I have achieved recognition as a public [speaker](#) and [teacher](#).
- o I enjoy **mentoring** people that want to start a career as a Data Scientist, both internally and in [external events](#), for which I have also gained [recognition](#).

Charles River Associates

London, UK
August 2018 – July 2019

Consulting Associate

Competition Practice – Working in mergers and abuse of dominance cases.

- o Successfully advised PayPal in its acquisition of iZettle, a company in the card payments industry, in a $2.2bn operation.
- o Worked on other confidential cases in the tech, air transport and natural gas sectors in front of the UK's and the EU's competition authorities.
- o Responsibilities ranged from analysing large datasets in **Python** and **Stata** to drafting economic reports to be presented as evidence in my cases.

Neovantas Consulting

Madrid, Spain
January 2016 – August 2017

Business Analyst – Data Analyst

- o **Banking sector** – Led a lighthouse project for the design and implementation of debt recovery strategy for a collection agency using speech analytics and econometric techniques. Improvement of up to +29 p.p. in collection ratios. Started this project as an intern; after one year, I was leading a team of three people, including client-facing responsibilities.
- o **Banking sector** – Built for major European bank building an econometric model to evaluate the impact of internal management training.
- o **Mentor** of new colleagues.

Education

Barcelona Graduate School of Economics

2017 – 2018

MSc in Economics and Finance - 8.8/10

Specialised in Econometrics and Machine Learning

Universidad Carlos III de Madrid

2012 – 2016

BA in Economics - 8.2/10

University of California, Berkeley

2014 – 2015

BA in Economics – Exchange Year (3) – 8.5/10

Languages

- o **Spanish** – Native
- o **English** – Full proficiency
- o **French** – Intermediate

Interests

I love scientific writing. I have published several blogs in Towards Data Science and various IBM sites, on topics such as [language models for cuneiform texts](#), [Bayesian networks](#) and more. You can find all my blogs [here](#). I also have a [podcast](#) (in Spanish) about ancient literature, my other passion.